REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
14-09-2000

2. REPORT DATE
September 14, 2000

3. DATES COVERED *(From - To)*
March 1995 - Sep 2000

4. TITLE AND SUBTITLE
Automated Synthesis of Multi-Agent Control

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-98-1-0130
N00014-95-1-0759

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)
Maja J. Mataric'

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Brandeis University
University of Southern California

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We have been pursuing a synthetic approach to studying the problem of controlling complex multi-robot systems by simultaneously developing a theory and testing it on complex domains consisting physical mobile robots. This process allows us to evaluate, improve, and further develop our theory, while producing a set of useful software and hardware applications. Our approach is behavior-based; the robots use a set of behaviors (parametric, goal-achieving control laws) as a substrate for control, representation, and learning. This approach scales well to large multi-robot systems, and enables us to flexibly explore complex problems such as the coordination of decentralized groups and learning in such distributed systems.

15. SUBJECT TERMS
Multi-robot control, robot learning, synthesis and analysis of robot behavior, mobile robots, behavior-based robotics

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
Unclassified	Unclassified	Unclassified	Unclassified	30	Prof. Maja J. Mataric'
					19b. TELEPONE NUMBER *(Include area code)* (213) 740-4520

Standard Form 298 (Rev. 8-98)

ONR Final Report
Grant title: **Automated Synthesis of Multi-Agent Control**

Principal investigator: Maja J Matarić
Computer Science Department
University of Southern California
941 West 37th Place, SAL 228, Mailcode 0781
Los Angeles, CA 90089-0781
tel: (213) 740-4520, fax: (213) 740-7285
mataric@cs.usc.edu
Grant number: N00014-95-1-0759
End date: Aug 31, 2000 (includes one year no-cost extension)

Abstract

We have been pursuing a synthetic approach to studying the problem of controlling complex multi-robot systems by simultaneously developing a theory and testing it on complex domains consisting physical mobile robots. This process allows us to evaluate, improve, and further develop our theory, while producing a set of useful software and hardware applications. Our approach is behavior-based; the robots use a set of behaviors (parametric, goal-achieving control laws) as a substrate for control, representation, and learning. This approach scales well to large multi-robot systems, and enables us to flexibly explore complex problems such as the coordination of decentralized groups and learning in such distributed systems.

1 Goals of the Project

Our goal in the work funded by ONR and reported here was to provide techniques that facilitate the development of effective multi-robot control systems that are:

- **Robust** to individual robot failures and communications failures;
- **Adaptive** to environmental and system changes;
- **Efficient** with respect to computation, communication, hardware requirements, energy expenditure, and environmental resources.

The techniques we developed succeed in meeting all of the above goals. They are principled and generally-applicable, leading to easily modifiable and analyzable systems. Furthermore, it is essential that these techniques were validated using groups of physical mobile robots in a variety of task domains.

2 Overview of Research

A large body of work was produced under the funding provided by this grant. It can best be summarized in the following three contribution categories:

1. Methods for robust, efficient **distributed behavior-based control of robot teams:** i) basis behaviors and ii) port-arbitrated behavior coordination;

2. Methods for **on-line real-time modeling of interaction dynamics**, using the underlying behavior-based control structure;

3. A large set of **validated coordinated multi-robot systems** demonstrating: chaining, robot soccer, variations on foraging (collection & coverage), and multi-target tracking.

This report describes each of these contributions, briefly describing the approaches, and providing a complete list of publications (by topic as well as cumulatively) and associated project Web sites, for additional information.

The rest of the report is structured as follows. We first describe our research into *behavior-based control of multi-robot teams* which addressed the issues of robustness, adaptivity, and efficiency. Next we describe *broadcast of local eligibility*, a general method we developed for coordinating collections of robots, based on well-defined port-arbitrated behavior messaging. Next we describe the methodology we developed for on-line real-time statistical modeling of interaction dynamics through using *augmented Markov models*, founded on the well-understood theory of semi-Markov chains. We conclude the report with a list of specific scientific and Navy/Dod contributions and the complete publications list.

3 Behavior-Based Control of Multi-Robot Teams

While terminology and some concepts of behavior-based robotics have become widespread, the central ideas are often lost as researchers try to scale behavior to higher levels of complexity. "Hybrid systems" which deliberate plans in terms of behaviors rather than simple actions have become common for higher-level behavior. Our research has demonstrated that a strict behavior-based approach can scale to higher levels of complexity than many robotics researchers assume, and that the resulting systems are in many cases more efficient and robust than those that rely on "classical AI" deliberative approaches. Our focus is on systems of cooperative autonomous robots in dynamic environments.

Though widespread in use, the term "behavior-based" lacks a clear, exact definition. Matarić (1997) gives an overview of common conceptions of the behavior-based approach. Brooks (1991*a*) describes a set of four key concepts essential to behavior-based robotics: *situatedness* - the use of the world as its own best model, *embodiment* - use of the world to ground regress, *intelligence* - as determined by the dynamics of interaction with the world, and *emergence* - intelligence as behavior in the eye of the beholder. Behavior-based systems thus are structured in terms of the observable activity that they produce, rather than traditional functional decompositions (Brooks 1991*b*). The activity-producing components, *behaviors*, compete for actuator resources and share perceptions of the world rather than any centralized representation. Behaviors tend to be simple, so that computational "depth" – the amount of computation that takes place between sensory perception and actuator commands – is minimized to maintain a high degree of interactivity with the environment. Behavior-based systems are highly parallel so that capability – new behaviors - can be added as increased computational "breadth." Behaviors are "layered" in such a way that capability is incrementally added to a functional system, leading to a design process that goes not from isolated components to a final system which integrates them into meaningful behavior, but from simple yet complete behavior to more complex complete behavior (Brooks 1991*b*, Brooks 1990*a*, Matarić 1995a). The design of behavior-based systems is thus often referred to as a "bottom-up" process (Brooks 1990*b*, Steels 1994), but this refers not so much to determination of the structure of the system as to a basis in physical sensing and action, and incremental development of sophistication from simple to complex. The system structure undergoes drastic changes driven by top-down task constraints as well as bottom-up sensorimotor constraints until a set of basis behaviors is determined (Matarić 1995); it is only with this solid foundation that the design process becomes one mainly of synthesis.

Basis behaviors (Matarić 1995) are a set of minimal behaviors that are sufficient to be combined into solutions to a class of tasks. Our early research (Matarić 1995a) on group behavior showed how various complex, biologically-inspired group behaviors could be composed from a set of general basis behaviors for spatial tasks, through two operators, *summation* of outputs and *switching* of outputs. *Flocking*, for example, is achieved by the summation of *homing, dispersion, aggregation,* and *safe-wandering*, while *foraging* results from switching (based on sensory conditions) between *safe-wandering, dispersion, homing,* and *following*.

The choice of basis behaviors has great influence on the efficiency of both the development process and the final system. Effort expended in refining basis behavior choices is usually paid back many times over; it is all too easy to reach (and sometimes difficult to detect) a state where a good percentage of a system's code is dedicated to working around earlier implementation choices. A good set of well-defined basis behaviors form a highly-reusable library of code; only a small amount of coding (if any) need be done to add "higher layers" which perform new tasks.

The following is a list of PI's publications that define, explain, and survey behavior-based control.

Published Papers:

Matarić, Maja J., "Coordination and Learning in Multi-Robot Systems", IEEE Intelligent Systems, Mar/Apr 1998, 6-8.

Matarić, Maja J., "Behavior-Based Robotics as a Tool for Synthesis of Artificial Behavior and Analysis of Natural Behavior", *Trends in Cognitive Science*, 2(3), Mar 1998, 82-87.

Matarić, Maja J., "Behavior-Based Control: Examples from Navigation, Learning, and Group Behavior", *Journal of Theoretical and Experimental Artificial Intelligence*, special issue on Software Architectures for Physical Agents, 9(2-3), H. Hexmoor, I. Horswill, and D.

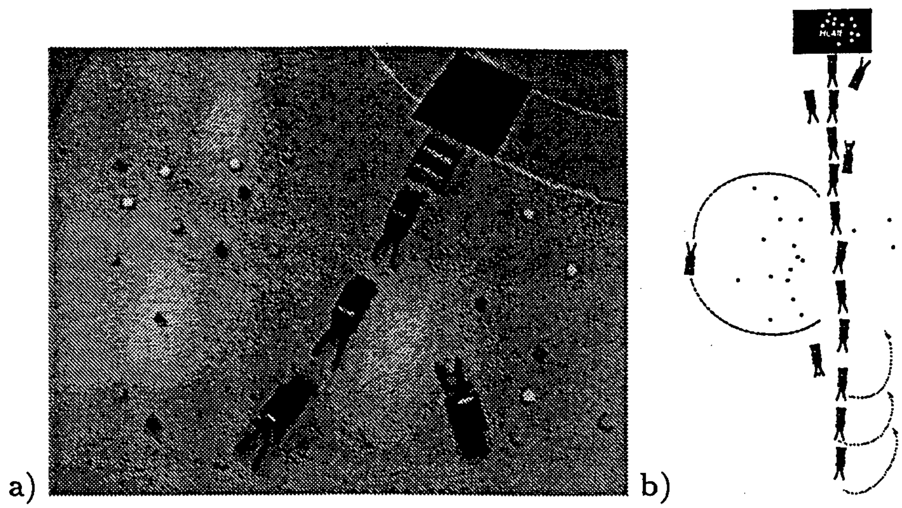


Figure 1: Foraging with a robot chain. **a)** A robot returns to the chain carrying a puck after a circular excursion. **b)** The robot at the end of the chain leaves to become a forager if it notices that many successful foragers are coming along the chain (indicating that the chain has grown past a rich deposit).

Kortenkamp, eds., 1997, 323-336.

Matarić, Maja J., "Behavior-Based Robotics", invited contribution to the *MIT Encyclopedia of Cognitive Science*, R. Wilson and F. Keil, eds., MIT Press, April 1999, 74-77.

The PI also maintains several Web informative pages on the topic:
http://robotics.usc.edu/~maja/robot-control.html
http://robotics.usc.edu/~maja/bbs.html
http://robotics.usc.edu/~maja/gruop.html

Next we present three research projects which examine three types of behavior-based coordination of multi-robot systems. All three are inspired to some degree by biological systems: one strives to recreate specific navigational techniques of ants, one uses different types of arbitration borrowed from various natural systems in a foraging task, and the third applies principles of environmental interaction abstracted from natural systems to teamwork in Robot Soccer. Further, all three relate deeply to the concept of situatedness discussed above – interaction through the environment. The robot chaining and soccer systems take active advantage of physical interactions between robots, while the different arbitration schemes for foraging are analyzed with respect to their ability to prevent destructive interactions (i.e., physical interference, or collisions).

4 Robot Chaining

Our *robot chaining* research was performed as an attempt to reproduce the stigmergic techniques[1] and benefits of pheromone-trail formation by ants. In the natural systems, individual ants deliberately encode information into the physical environment (by depositing chemicals known as pheromones), and over time interesting global properties emerge that allow these chemical markings to be used as a navigational aid for position-dependent tasks. The release of pheromones leads to trails that can be followed, which are subject to decay of pheromone strength over time. When pheromones are released only during certain phases of a task (e.g., while carrying some item back to the nest), trails can begin to form efficient paths to useful locations, such as rich supply areas. Since paths that take less time to traverse (and are thus traversed more frequently) gain more pheromone strength than longer ones, a very simple control strategy of probabilistically choosing the "strongest" path leads to group behavior that adjusts to follow dynamically determined shortest paths to dynamically changing useful destinations.

Our robot chaining system for foraging replaces the chemical pheromones of the ant trails with the physical bodies of simple robots (as illustrated in Figure 1). We have demonstrated that a group of robots equipped with only physical contact sensors is able to form a physical pathway that members of the group can use for navigation. The behavior of *chain-following* consists of moving in arcs that guarantee intermittent contact with the chain (much as we might guide ourselves by tapping a hand against a wall in the dark). The *search* behavior is performed through "circular excursions," in which the robots hold a (random) steady steering angle so as to explore an area next to the chain while being able to regain contact with the chain without need for odometry or other non-contact sensors. A *join-chain* behavior can be used as robots reach the end of the chain, through a protocol of taps exchanged by the current "last link" and the robot attempting to join. The robots that are part of the chain maintain chain integrity through a *link* behavior intermittent contacts, using a similar tapping protocol.

Since the links of the chain are capable of computation and motion, rather than depositing pheromones and having paths "emerge" through chemical processes the chain links can collect some statistics of the activity of the chain-following robots, and use them to adapt to the environment by physically modifying the chain. Two types of chain modification are sufficient for generating an optimal path to a rich source in a plane with no insurmountable obstacle: shifting of chain direction, and lengthening/shortening of the chain.

Natural ants change roles (e.g., from foragers to internal nest workers) in response to the number of encounters each ant has with ants fulfilling *other* roles – a nest worker that encounters a number of successful foragers in a given time period will decide to forage. As seen below, the process we describe for adjusting the length of the chain functions in a very similar manner.

In order for the chain to move to intercept a rich source, all that is necessary is for the chain links to monitor how many times they have had Success Reports on their right and left sides. If basic behaviors are in place that maintain chain integrity, individual robots can

[1] Stigmergy refers to the various means of interaction through the environment rather than through direct communication.

shift towards the direction of more Success Reports (within constraints of chain integrity) without need for explicit communication with neighboring links. In this way, the entire chain will slowly shift towards a rich source.

In order to more clearly replicate the ant systems, and eliminate the risk of the chain infinitely extending in a direction with no sources, it would be necessary to introduce random direction-shifting of chain links with some probability. Decay of trails could be replicated in two ways: either the links could factor recency into their statistics, or, more minimally, the links could merely react by shifting towards the direction of every Success Report, allowing such temporally-based statistics to to be computed "physically."

Ideally, once the chain has shifted to intersect a rich source, we would like it to end there – that is, we would like the end of the chain to be near the center of the richest area, so that robots can return directly from the source to Home. In situation where the chain extends past a rich deposit, the chain should be shortened in order to both optimize the pathway and allow more robots to participate in transport of material.

There are two ways for this to happen; in either case, the chain will tend to shorten to the optimal length when there is a rich deposit, and naturally begin to grow again if this source begins to be exhausted. One way is for the chain links to collect Success Report statistics (most likely, the number of recent Success Reports at each link, for comparison) and pass them along the chain through some protocol, allowing the end-of-chain robot to decide when it should leave the chain and become a forager (by passing end-of-chain status to the preceding link).

A more minimal, situated way to adjust the chain length is to simply have the end-of-chain robot leave the chain after a period of time. If the chain extends past a rich source, there will be fewer robots attempting to append themselves to the end of the chain (since many will be carrying material and thus be ineligible); if the chain does not reach a source, few if any robots will be carrying and thus most will attempt to append themselves and lengthen the chain. This can be seen as dynamic role assumption such as (Gordon 1999) finds in ant colonies: when the end-of-chain encounters mostly successful foragers (which do not attempt to append themselves to the chain), it is likely to leave the chain and become a forager. When the foragers encounter mostly chain links without finding useful material, they tend to become chain links. The robots, like the ants, fulfill roles as determined by global constraints.

Through a physically-situated approach, robots are able divide themselves efficiently into foragers and chain links and perform position-dependent tasks using only local sensing and interaction. Werger & Matarić (2000) discusses further interesting properties of the chaining system regarding efficient role assumption given the inherent physical heterogeneity of the particular robots used.

Published Papers:

Werger, Barry B. and Matarić, Maja J., "Robotic Food Chains: Externalization of State and Program for Minimal-Agent Foraging," *From Animals to Animats 4, Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, MIT Press, pp. 625-634.

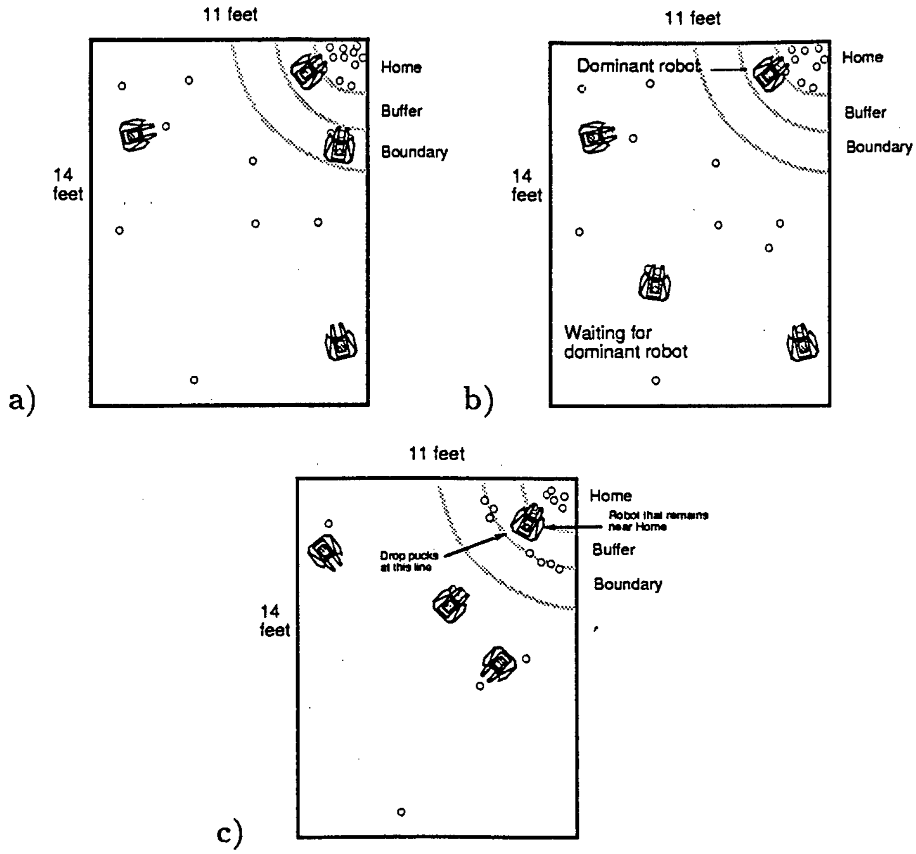


Figure 2: Three versions of the foraging task. **a)** Homogeneous: all robots are behaviorally identical and act independently. **b)** Pack: robots are organized in a dominance hierarchy. **c)** Caste: robots are behaviorally differentiated and occupy different regions of the task space.

Werger, Barry B. and Matarić, Maja J. "Exploiting embodiment in multi-robot teams", USC Institute for Robotics and Intelligent Systems Technical Report IRIS-99-378, 1999.

More information about this work can be found on the project web page: **http://robotics.usc.edu/~barry/Chaining.html.**

5 Ethologically-Inspired Foraging

Social structure plays an important role in the performance of a group, whether it consist of biological or synthetic individuals. In a synthetic approach, such as mobile robotics, it may is difficult to determine an appropriate social structure for a group performing a specific task. Issues to be considered include how many robots to use, and how the task should be divided both temporally and spatially among the individuals in order to allow completion of the task and provide a desired level of performance.

A pragmatic, principled approach to guide the resolution of these issues is desirable. We have explored such an approach based on the analysis and manipulation of physical *interference* (i.e., collisions) a readily measurable property of mobile robotic systems. Our approach involves a controller refinement methodology that is motivated by biological evolution and based on the application of ethologically inspired *arbitration schemes*, i.e., modifications to social structure, or the multi-robot controller.

In our approach, the first multi-robot controller that is constructed for a desire task is homogeneous, loosely analogous to the herd phenomenon exhibited by certain animal species. In such a controller, the robots are behaviorally identical, each capable of independently completing the entire task. Since the robots function independently of each other, there is no need for explicit communication. The homogeneous controller enables a base-case analysis of interference characteristics. This initial controller is refined by modifying its interference characteristics through the employment of *pack arbitration* or *caste arbitration.*

Pack arbitration is modeled after the phenomenon of the pack observed in wolf and other animal societies. In these, any individual is physically and behaviorally capable of performing most functions necessary to the group. In order to minimize aggressive behavior which, if not controlled, can jeopardize the pack, a form of dominance hierarchy exists among the individuals. Similar to animal packs, in pack arbitration, all of the individuals of the robot group are physically and behaviorally capable of performing any of the functions necessary for the group to complete the task (as is also true for the herd scheme). To avoid interference (collisions) between individuals, the controller is modified so that the robots take turns entering regions where interference was high in the homogeneous case, with the most dominant robot going first. This form of arbitration contains some implicit assumptions about communication. The robots must be able to communicate their rank and intention to enter a region of potentially high interference. In addition, they must be able to determine when a dominant robot has failed so as not to wait indefinitely for it to complete its objective.

Caste arbitration is modeled after the structure apparent in many social insect societies. In these, individuals are behaviorally heterogeneous and are not capable of accomplishing all of the tasks that the group requires. Individuals may also be physically differentiated. As an example, consider many ant species whose colonies include worker, drone, possibly warrior castes, and at least one queen. Each individual is a member of one of these *castes* and has associated physical and behavioral characteristics. No one caste can maintain the colony without the others.

In caste arbitration, physical interference between robots is modified through the use of territoriality, with different castes occupying different regions of the task space and potentially having different behavioral repertoires. This limits destructive interactions such as collisions. Robustness in caste arbitration is achieved by allowing members to change castes when necessary. If, for example, all the members of one caste fail, a member of some other caste must be able to take over. Some form of communication is needed to determine the number (or density) of individuals in each caste. Such caste switching is observed in honey-bee societies (McFarland 1987).

We have demonstrated our interference-modifying approach to controller refinement by implementing homogeneous, pack, caste, and territorial behavior-based controllers for a foraging (object collection) task, a prototype for various applications including distributed solutions to de-mining, toxic waste clean-up, and terrain mapping (Figure 2). The experiments required four physical mobile robots to search an 11×14 foot region for pucks and bring them to a designated goal location. We evaluated and compared the controllers according to three performance criteria: time-to-completion, inter-robot collisions (interference), and energy expenditure. An important component of this analysis was the comparison of internal behavior activations to the externally observed interference. This initial study of behavior activations inspired our later efforts in modeling interaction dynamics using behavior activations and *augmented Markov models.* A parallel effort in our work on ethologically-inspired foraging aimed at demonstrating the ease with which robust, easily modifiable behavior-based controllers may be designed, implemented, and evaluated.

Published Papers:

Goldberg, Dani and Matarić, Maja J., "Design and Evaluation of Robust behavior-Based Controllers for Distributed Multi-Robot Collection Tasks", in "Robot Teams: From Diversity to Polymorphism", Tucker Balch and Lynne E. Parker, eds., 2001.

Goldberg, Dani and Matarić, Maja J., "Robust Behavior-Based Control for Distributed Multi-Robot Collection Tasks", USC Institute for Robotics and Intelligent Systems Technical Report IRIS-00-387, 2000. Also submitted to *IEEE Transactions on Robotics and Automation.*

Fontan, Miguel S. and Matarić, Maja J., "Territorial Multi-Robot Task Division", *IEEE Transactions on Robotics and Automation*, 14(5), Oct 1998.

Goldberg, Dani and Matarić, Maja J., "Interference as a Tool for Designing and Evaluating Multi-Robot Controllers", *Proceedings of the Fourteenth National Conference on Artificial Intelligence*, AAAI Press, 1997.

Goldberg, Dani and Matarić, Maja J., "Interference as a Guide for Designing Efficient Group Behaviors", Brandeis University Computer Science Technical Report CS-96-186, 1996.

Fontan, Miguel S. and Matarić, Maja J., "A Study of Territoriality: The Role of Critical Mass in Adaptive Task Division", *Proceedings, From Animals to Animats 4, 4th International Conference on Simulation of Adaptive Behavior (SAB-96)*, P. Maes, M. Matarić, J-A. Meyer, J. Pollack, and S. Wilson, eds., MIT Press, 1996, 553-561.

More information about this work can be found on the project Web page:
http://robotics.usc.edu/ dani/hetero-homogeneous-groups.html.

5.1 Minimalist Robot Soccer

Robot soccer has become the recognized benchmark challenge domain for both mobile robotics and Artificial Intelligence in general. Because the task requires both real-time tactics and higher-level strategy, in a context that involves both cooperation (within a team) and competition (between teams), it presents a set of challenges that is uniquely complex, and thus progress on this problem has implications in various application areas. While robot

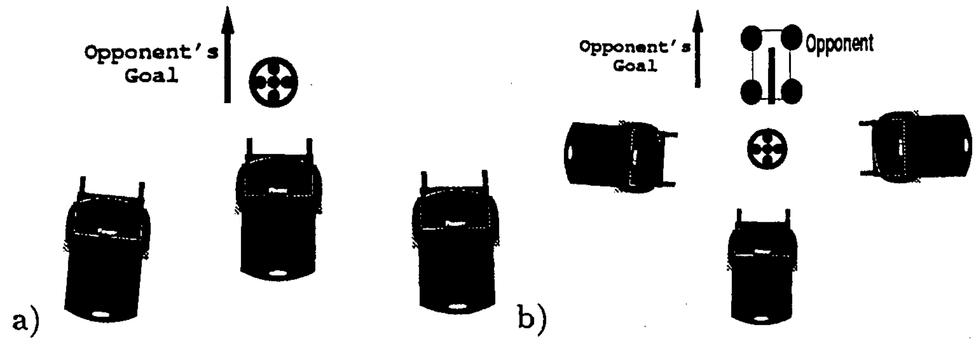


Figure 3: Formations in Robot Soccer. **a)** Offensive: Interaction of simple behaviors causes the robots to fall into a V-formation when the ball is in motion roughly towards the opponent's goal. Perceptual properties limit the formation to three robots. **b)** Defensive: When the ball is not moving roughly towards the opponent's goal, the robots cluster around it to form an effective barrier and be in good positions for recovery.

soccer has not been a major area of our research, we have successfully validated our behavior-based methodology in this domain. We used minimalist behavior-based techniques to design a simple control systems that displays highly sophisticated individual and team behavior including effective obstacle avoidance in a dynamic environment, generation of smooth, effective trajectories, three separate methods of ball handling, and dynamic configuration into appropriate population-limited offensive and defensive formations. The robots use no explicit communication, and for the formations, they are able to use local interactions to determine globally optimal roles.

Werger (1999) discusses at length our minimalist approach to team cooperation for a robot soccer team. Though individual players can perceive only the ball, the goals, and obstacles (which are not distinguished but may be walls, opponents, or teammates), and have no communication equipment, the team displays sophisticated cooperative behavior. The team falls into appropriate formations for offensive and defensive situations with the interesting property of formation size limitation.

The cooperative behaviors result from the interaction of simple individual behaviors. *Push* causes the robot to line up behind the ball and push it towards the opponent's goal. A second behavior, *Safety*, causes the robot to maintain the maximum safe velocity (as determined by sonar sensors). A third behavior, *Disperse*, causes the robot to rotate away from anything too close to its sides. Finally, a *Patrol* behavior causes the robot to patrol its half of the field defensively when it has not perceived the ball for a few seconds.

In an offensive situation, seen in Figure 3a, one robot serendipitously gets to the ball first and begins to *Push* it forward. Teammates also try to *Push*, but their *Disperse* and *Safety* behaviors slow them down and steer them away when they get very close to the *Pushing* robot, and thus tend to fall into a V-formation.

This formation provides effective "fumble protection" that is essential in the robot soccer domain. Robots often accidentally knock the ball off course while dribbling it forward; this formation provides backup and recovery. With this formation it is not uncommon for possession of the ball to transfer between the robots of an advancing group without loss of possession by the team. The formation also provides for a very quick defense if the ball is stolen (see below).

The size of the offensive formation is limited by the interaction between the four behaviors above and the physical bodies of the robots. Once there are three robots in the formation, any other robot trying to *Push* the ball will have its view of the ball occluded by the bodies of the first three robots in the formation. When this occlusion lasts for more than a second or two, the *Patrol* behavior gains control of the robot and it gives up on following the ball. In this way, necessary roles are filled (attacker, supporters, and defense) without negotiation, explicit definition or assignment of roles, or even any representation of teammates.

In a defensive situation (as in Figure 3b) the ball is not advancing toward the opponent's goal. The same behaviors described above cause the robots to fall into a semi-circular arrangement around the ball rather than the V-formation of the advance, since the robots on the sides are no longer kept behind by lower speed. This formation very effectively prevents the opponent from continuing to move the ball up the field, and places players in a good position to gain possession of the ball. An emergent "batting behavior" (another result of the interaction between the four behaviors listed above, described in Werger (1999)) makes it likely that the *Push*ing robot will jostle the ball towards one of its teammates, which can smoothly begin an advance from the side; this can be seen as a rudimentary form of ball-passing.

Transition between offensive and defensive formations is determined by motion of the ball, and is not even perceived by the robots; there is no concept of "offensive" or "defensive" (or even of "formation") anywhere in the behavior structure. Simple sensing of the local environment leads to flexible, dynamic team behavior that many researchers claim requires higher deliberation and explicit communication.

Thus, in our soccer system, the situated approach allows robots to efficiently assume roles in offensive and defensive formations as determined purely by physics-inspired interaction and visual occlusion. Simple, stateless control allows sophisticated behavior including dynamically-determined limited-size formations, maintenance and recovery of ball possession, and simple passing. Assumption of roles takes place without any communication or explicit representation or coding of roles – the role behavior "emerges" from the interaction of a few simple behaviors.

Published Papers:

Werger, Barry B. "Cooperation Without Deliberation: A Minimal Behavior-based Approach to Multi-robot Teams", *Artificial Intelligence*, 110, 1999, 293-320.

Minoru Asada, Peter Stone, Hiroaki Kitano, Barry B. Werger, Yasuo Kuniyoshi, Alexis Drogoul, Dominique Duhaut, Manuela Veloso, Hajime Asama and Sho'ji Suzuki, "The RoboCup Physical Agent Challenge: Phase I". Applied Artificial Intelligence (AAI), Volume 12, 1998.

Barry B. Werger, "The Spirit of Bolivia: Complex Behavior Through Minimal Control", in Proceedings of RoboCup 97, Nagoya, Japan, 1997.

Barry B. Werger, "Principles of Minimal Control for Comprehensive Team Behavior", Pro-

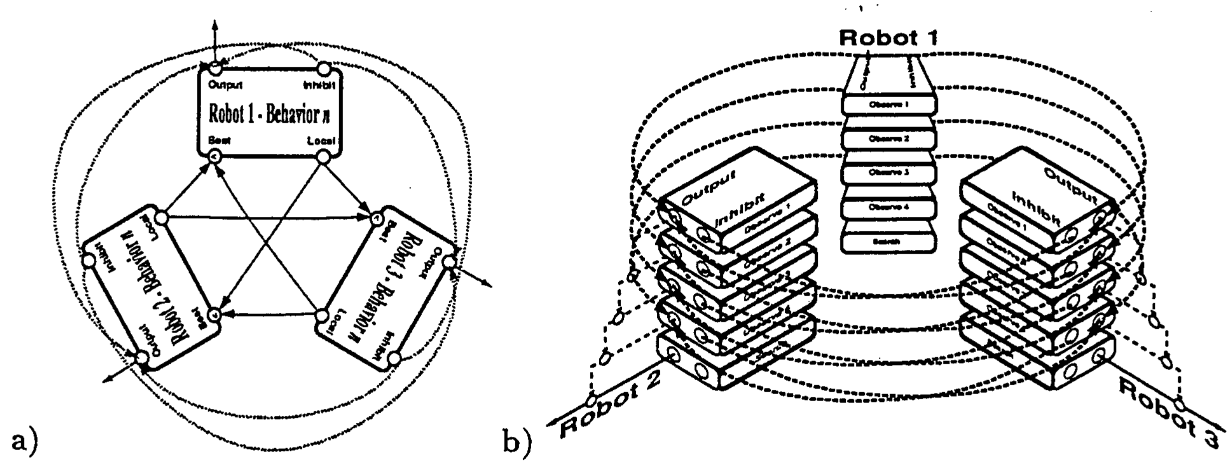


Figure 4: **a)** *Cross-Inhibition*: A cross-inhibited peer group. The Local port of each robot's behavior B_n broadcasts a locally-computed eligibility estimate to the Best port of each other robot's behavior B_n. Each Best port maintains the maximum of the eligibility messages it has received in the current decision cycle. Whichever robot has a local eligibility better than or equal to the Best it receives writes to its Inhibit port, causing write-inhibition of behavior B_n's Output port(s) in the other robots, thereby "claiming" the task. **b)** *Cross-Subsumption*: The structure of a cross-subsumptive system. Subsumption is used to arbitrate within each robot between cross-inhibited behaviors. Some lines are omitted for clarity; each "layer" is connected as in **a)**.

ceedings of ICRA-98.

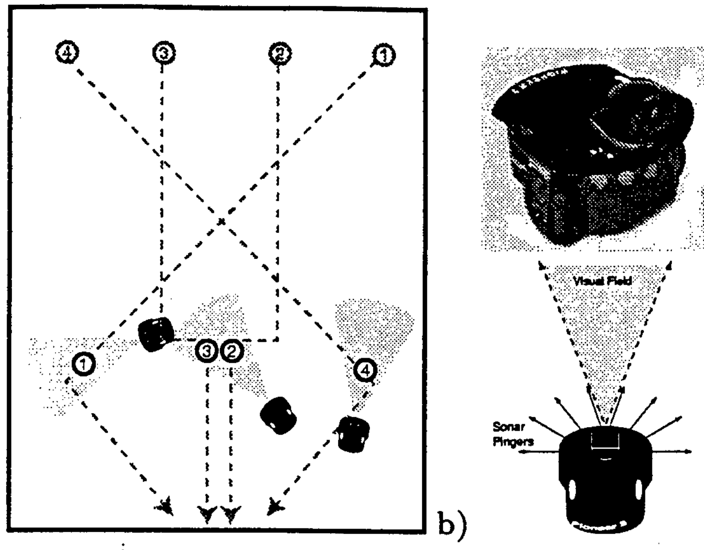
Barry B. Werger and Maja J Matarić, "Quick'n'Dirty Generalization for Mobile Robot Learning" presented as a poster at IJCAI-97.

Barry B. Werger, "Multiple Agents From the Bottom Up", in Proceedings, Fourteenth National Conference on Artificial Intelligence (AAAI-97), Providence, RI, 1997.

More information about this work can be found on the project web page:
http://robotics.usc.edu/~barry/ullanta/UPRsoccer.html.

6 Broadcast of Local Eligibility for Group Coordination

Our Broadcast of Local Eligibility project investigates the possibilities of extending the *port-arbitrated behavior* (PAB) paradigm across networks of robots. While it has often been hypothesized that there need be no distinction between inter-robot and inter-behavior communication, no previous system has provided standard tools that allow port-based messaging, suppression, and inhibition between behaviors on separate networked robots. Our intention



a) b)

Figure 5: a) *CMOMMT Experiments*: CMOMMT experiments require a team of three robots to maintain continuous observation of four moving targets in an 18 by 22 foot enclosure. Robots are shown with observation ranges; fields of view extend further. Targets are numbered circles. Light grey targets and dashed lines indicate initial positions and paths of targets. b) *Robot Testbed*: Three Pioneer 2DX robots.

is to demonstrate that behavior-based systems restricted to well-defined port-arbitrated interactions can scale to higher levels of competence than is generally assumed. Specifically, we show that when the port-arbitration paradigm is extended across networks, the resulting systems are able to dynamically reconfigure themselves in order to allocate resources in response to task constraints, environmental conditions, and system resources. We have developed the Broadcast of Local Eligibility (BLE) as a general tool for coordination between robots.

In port-arbitrated behavior-based control (PAB) systems, controllers are written in terms of behaviors, which are groups of concurrent processes that share a public interface. This interface is composed of ports, which are registers that each hold a single data item (e.g., an integer, oat, string, or complex data structure).

Ports in different behaviors are linked together by connections, which are unidirectional data paths between a source port and a destination port. A port can have any number of incoming and outgoing connections. When data is written to a port, either directly from a process within the behavior or indirectly through a connection, it is generally propagated along all of that port's outgoing connections. We say "generally" because data flow can be modified by special connections which may suppress, inhibit, or override data flowing through other connections.

It is through these mechanisms of suppression and inhibition that subsumption hierarchies, as well as other forms of arbitration, can be efficiently and intuitively implemented. Since connections are external to the behaviors, behavior code is easily re-usable, and interaction between behaviors can be modified dynamically. The port abstraction enforces a data-driven approach to programming that "grounds" computation in sensor readings and effector actions. The PAB approach allows a clean, uniform interface between system components (behaviors) at all levels that abstracts away many issues of timing and communication; the "black boxes" of behaviors may contain reactive mappings or deliberative planners. While our research focuses on non-deliberative approaches, we believe that PAB interaction between system components can help reduce the complexity of the components themselves, whatever their type.

Our Broadcast of Local Eligibility (BLE) mechanism, illustrated in Figure 4, is a standard tool comprised of three specific ports added to BLE-arbitrated behaviors – *Local*, *Best*, and *Inhibit*. Each robot makes a local (i.e., derived from data from the robots own sensors) estimate of its own eligibility for a some task. This eligibility estimate is written to the appropriate behavior's *Local* port, which is connected so as to broadcast this estimate to the *Best* port of each behavior of the same name on every robot on the local network. The *Best* port filters all the incoming messages for the maximum. A comparison is made between the locally determined eligibility (the *Local* port's value) and the best eligibility calculated by a peer behavior on another robot (the *Best* port's value). When a robot's local eligibility is best for some behavior B_n which performs task T_n, it writes to it's *Inhibit* port, which is connected so as to inhibit the peer behaviors (that is, behaviors B_n) on all other robots. In this manner, the most eligible robot "claims" task T_n. Since this inhibition is an active process, failure of a robot which has claimed a task results in the task being immediately "freed" for potential takeover by another robot. Since BLE is based on broadcast messages and receiving ports that filter their input for the "best" eligibility, BLE-based systems are inherently scalable. Up to the limit of communication bandwidth, any number of BLE-enabled robots added to a system will properly interact. BLE allows heterogeneous robots to efficiently allocate themselves to appropriate tasks without the need for any explicit communication or global knowledge of particular abilities. The ability to dynamically instantiate and connect BLE-enabled behaviors allows systems to scale in capability as well as in number of robots.

We have validated our BLE approach through experiments in the domain of cooperative multi-robot observation of multiple moving targets, or CMOMMT. CMOMMT involves a team of robots which must attempt to keep a number of prioritized moving targets under constant observation (as illustrated in Figure 5. To do this, each robot has behaviors referred to as *Observers*, each of which is parameterized to cause the robot to attempt to stay within observation range of a specific Target (i.e., *Observer1* causes a robot to track *Target1*). A *Search* behavior on each robot causes the robot to wander randomly (intended to be used when no suitable Targets are within the visual field). BLE was used to arbitrate between these behaviors, that is, to determine which task (a specific target or search) each robot in the system should attend to. Results have demonstrated that BLE is able to efficiently assign robots to subtasks in response to differences in robot capabilities and environmental situations, maintaining better coverage of targets than three other arbitration schemes used for comparison.

Scientifically, our research of the Broadcast of Eligibility (BLE) technique demonstrates that the port-arbitrated behavior-based control paradigm (PAB) can be extended in such a way that robust, scalable, fully-distributed control for robot teams can be designed and implemented in a principled manner. A standardized, general technique such as BLE is a major step towards rigorously-analyzable behavior-based systems; lack of analytic techniques has often been pointed to as a weakness of behavior-based systems.

Further, we have demonstrated that PAB interaction, and BLE in particular, are principled means of gaining many of the advantages of biologically-inspired, situated systems. Previous insect-inspired multi-robot systems were able to take advantage of the fact that they were situated in the physical world to gain robustness and scalability while minimizing requirements for local (individual-robot) complexity, but these systems were constructed in a fairly ad-hoc manner. Our research has shown that PAB systems can be seen as situated in an abstract "behavior space," and that BLE is able to structure this behavior space in a principled manner. BLE systems are as a result responsive to both their physical and behavior-space environments, gaining the benefits of situatedness while being quick and straightforward to design, implement, and analyze.

Practically, our work has demonstrated the abovementioned benefits, as well as the effectiveness of the resulting systems, in a multi-robot, multiple-moving-target observation task. Experimentation has shown BLE systems to be adaptive to unforeseen individual differences between robots as well as changing environmental situation and task coverage.

The PAB paradigm and BLE are based on "unreliable messaging" in which receipt of a sent message is never guaranteed. Systems are thus naturally designed to be robust to many types of communication failure, able to adapt automatically to variations in information and resource availability. This is particularly important underwater and surf-zone operations where communication bandwidth and availability are low.

The layered-behavior approach inherent in BLE allows "bottom-up" design of systems, in which simple individual behaviors can be well tested and then augmented with higher-level, but equally simple behaviors, which in turn can be thoroughly tested. Much as local interactions of ants that follow simple rules lead to complex, globally optimal activity, interaction of extremely simple behaviors both within and between robots lead to efficient global task assignment and performance. The practical benefits of the bottom-up approach have been widely demonstrated by the effects of the behavior-based "revolution" in robot control, but have not previously been combined with a principled inter-robot arbitration technique such as BLE. Thus, systems can be rapidly designed, component behaviors can be easily and widely reused, and systems can be incrementally tested. For both military and commercial applications, re-usability of well-tested components is of great benefit during design, deployment, and maintenance stages.

Finally, PAB and BLE have lead naturally to a change in the concept of a "control language" used to command and interact with multi-robot systems. Rather than a conventional Control Language which has basic commands related to vehicle capabilities (e.g., MOVE, REPORT), and the associated difficulties of uniform syntax and semantics across robots, our concept of a behavior-based control language with basic commands for behavior manipulation, and standard-interface behavior libraries. The language itself is thus both simplified and more flexible, and allows on-the-fly modification of system behavior in unforeseen ways. The ability to easily modify individual and group behavior at all levels, indeed to reconstruct controllers on-the-fly through simple, efficient behavior manipulations, speeds up the development process significantly and provides for emergency changes to deployed systems.

Published Papers:

Werger, Barry B. and Matarić, Maja J., "From Insect to Internet: Situated Control for Networked Robot Teams", *Annals of Mathematics and Artificial Intelligence*, 2000.

Werger, Barry B. and Matarić, Maja J., "Broadcast of Local Eligibility for Multi-Target Observation", *Distributed Autonomous Robotic Systems 4, Proceedings of DARS 2000*, Knoxville, Tennessee.

Duarte, Christiane N. and Werger, Barry B., "Defining a Common Control Language for Multiple Autonomous Vehicle Operations", *Proceedings of OCEANS 2000 MTS/IEEE*, Providence, Rhode Island, September, 2000.

Werger, Barry and Matarić, Maja J., "Broadcast of Local Eligibility: Behavior-Based Control for Strongly-Cooperative Robot Teams", *Proceedings of the Fourth International Conference on Autonomous Agents*, Charles Sierra, Maria Gini, and Jeffrey S. Rosenschein, eds., ACM Press, 2000, 21-22.

More information about this work can be found on the project Web page: **http://robotics.usc.edu/~barry/BLE**.

7 On-Line Modeling of Robot Interaction Dynamics

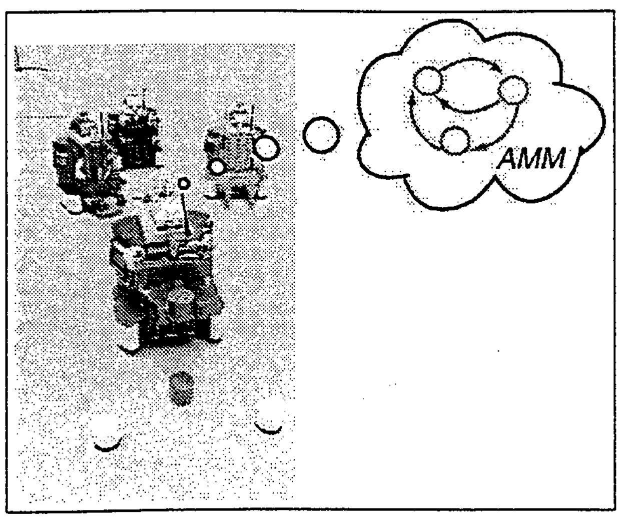


Figure 6: Each robot generates at least one AMM (depending on the application) during the execution of a task. The AMM captures statistics on behavior execution as the robot interacts with its environment.

Learning models of the environment, other robots, and interactions between them is a very challenging task in mobile robotics. Not only do noisy sensors and actuators pose inherent difficulty, but in the multi-robot domain which is the focus of our work, non-stationarity is an additional challenge. Limited computational and memory resources, in conjunction with often limited amounts of training data, can make brute force approaches to many learning techniques (i.e., hidden Markov models, partially observable Markov decision processes, reinforcement learning) intractable on mobile robots. In order to achieve convergence, it is often necessary to bias the learning system, for example, by providing an appropriate initialization, choosing a tractable search space, and/or making heuristic modifications to the learning algorithm.

We began our research by selecting the underlying behavior structure of the robot controller as the representational level for learning. In our first approach, we constructed trees of behaviors representing histories of their use (Michaud & Matarić 1997). This work was successfully demonstrated in the context of one or two concurrently learning mobile robots adapting to a changing environment (in some cases featuring a group of other, non-learning, unpredictable mobile roaming and interfering robots) in order to more efficiently perform a foraging task (specifically, finding a target object and delivering it to a goal location) (Michaud & Matarić 1998*c*). Besides being able to adapt their strategy to a non-stationary environment, the robots also demonstrated automated specialization: they adopted different but complementary strategies so as to minimize interference with each other (Michaud & Matarić 1998*a*, Michaud & Matarić 1998*b*).

In order to expand and generalize this idea of using behaviors themselves as the basis for a model, we then developed *augmented Markov models* (AMMs) as an approach to creating behavioral models of robot/environment interaction dynamics that accommodates the domain challenges and limitations mentioned above (Figure 6). The approach is computationally inexpensive, incrementally generating and modifying parsimonious models in real-time using only a small continuous stream of training data. For model generation, each data symbol indicates which behaviors in the robot are currently active. Because behaviors encompass both sensing and action, they provide a rich representational substrate for the models and help provide parsimony. In essence, an AMM is used as a repository of statistics about the execution of behaviors in a controller as the robot interacts with its environment and other robots. The basic structure of an AMM is a semi-Markov chain, with each state representing a behavior, and with probabilistic transitions (links) between states. The semi-Markov chain is augmented with statistics on state and link usage which are employed in model construction, modification, and utilization.

AMMs have a number of characteristics that make them naturally applicable to the robotics domain: they are compact, have a low computational overhead, and can be generated and used in real-time in one trial. We have demonstrated the applicability of AMMs to a number adaptation and learning problems in mobile robotics, described next.

Published Papers:

Goldberg, Dani and Matarić, Maja J., "Augmented Markov Models", USC Institute for Robotics and Intelligent Systems Technical Report IRIS-99-367, 1999.

Goldberg, Dani and Matarić, Maja J., "Mobile Robot Group Coordination Using a Model of Interaction Dynamics", *Proceedings of the SPIE: Sensor Fusion and Decentralized Control in Robotic Systems II*, Gerard T. McKee and Paul S. Schenker, eds., SPIE, 1999, 63-73.

Additional information about this work and the five AMM applications described in the following subsections can be found on the project web page:
http://robotics.usc.edu/~agents/projects/amms.html.

7.1 AMMs for Fault Detection

A robot's individual performance can impact the ability of a group to achieve effective group-level coordination. As an example, consider a scenario where a single robot develops a hardware failure and is neither able to complete its portion of the group task, nor to inform the other group members of its failure. If the members do not know to compensate for the incapacitated robot, the group as a whole may fail to complete its task. Monitoring individual robot performance, in this case for fault detection, is thus an important component of group coordination.

We limit our consideration of faults to those that would keep a robot in one behavior for an inordinate period of time. Such faults may include sensor and actuator failures, as well as the robot becoming physically stuck. To detect a potential fault, we used our AMM construction algorithm to compare, at each time step, the total time a robot has spent in the current AMM state to the mean and variance calculated from previous data for that state. A statistical confidence estimate on the upper bound of the mean is used to indicate a potential fault.

We tested this approach on-line by having the robots perform elements of a foraging task. If the model detected that the robot had been in one of the behaviors too long, it would send a signal to the robot, which would in turn beep (a call for assistance) to indicate a potential fault. We simulated a fault (the robot getting stuck on a rock) by lifting the drive wheels off the ground. During the dozen trials we conducted, the robot never failed to detect the fault.

Published Papers:

Goldberg, Dani and Matarić, Maja J., "Mobile Robot Group Coordination Using a Model of Interaction Dynamics", *Proceedings of the SPIE: Sensor Fusion and Decentralized Control in Robotic Systems II*, Gerard T. McKee and Paul S. Schenker, eds., SPIE, 1999, 63-73.

Goldberg, Dani and Matarić, Maja J., "Coordinating Mobile Robot Group Behavior Using a Model of Interaction Dynamics", *Proceedings of Autonomous Agents '99*, Oren Etzioni, Jorg P. Juller, and Jeffrey M. Bradshaw, eds., ACM Press, 1999, 100-107.

7.2 AMMs for Group Affiliation

The ability of a robot to determine what group it belongs to (i.e., its group affiliation) is another important component of group coordination. Suppose a robot were introduced into an environment containing several groups specializing in different tasks. In order to be able to coordinate its activity with the group it fits into best, it must have some mechanism for determining its group affiliation. In a learning system where the robot's final behavior is not predetermined, group affiliation is not designated *a priori.*

AMMs provide a mechanism for determining group affiliation. Two robots that wish to ascertain whether they belong to the same group can transmit data generated by their AMMs, then determine the probability of the other robot's data on their respective AMMs. If each AMM accepts data generated by the other's AMM (with probability >0), then the robots are designated as members of the same group. They are considered to have the same *ability*, or capacity for performing a particular task.

In addition to this coarse "don't accept"/"accept", or ability-based, determination of group affiliation, a more refined categorization can be made by considering the actual probabilities of symbol sequences. To test this notion, we conducted 2 sets of trials with the robots performing the wandering-avoiding behaviors. In one set of trials, the region was free of obstacles, in the other, it was sparsely distributed with small obstacles. Our hypothesis was that a data set from an AMM generated in one of the two environments should produce higher probabilities on the AMMs from that environment than on the ones from the other environment. This reliably proved to be the case after only a few minutes of model generation. These results, produced from little training data and very similar environments, suggest that AMMs can be used to make subtle behavioral distinctions. These distinctions can be thought of as experience-based. Since the robots are able to and do perform the same task, it is their specific individual experiences that differ, and are the basis for distinction.

Published Papers:

Goldberg, Dani and Matarić, Maja J., "Mobile Robot Group Coordination Using a Model of Interaction Dynamics", *Proceedings of the SPIE: Sensor Fusion and Decentralized Control in Robotic Systems II*, Gerard T. McKee and Paul S. Schenker, eds., SPIE, 1999, 63-73.

Goldberg, Dani and Matarić, Maja J., "Coordinating Mobile Robot Group Behavior Using a Model of Interaction Dynamics", *Proceedings of Autonomous Agents '99*, Oren Etzioni, Jorg P. Juller, and Jeffrey M. Bradshaw, eds., ACM Press, 1999, 100-107.

7.3 AMMs for Dynamic Leader Selection

Another issue impacting group coordination is performance. Consider a group of robots organized in a hierarchy. Due to inherent variations in sensors and actuators, or inexperience with a specific robotic platform, it may be difficult to accurately assess the ability of a robot at performing a novel task. Alternatively, even if performance history is available, there is no guarantee that future performance will neither improve nor degrade. The ability of individuals may change over time, but it is important that the performance of the group remain as high as possible. To achieve this, some mechanism for dynamic restructuring based on performance is necessary, especially in social structures such as hierarchies where significant reliance is placed on the most dominant individuals. We have explored dynamic leader selection using AMMs as one mechanism for restructuring hierarchies and maintaining or improving group performance.

In our experiments, four robots had to perform the foraging task, with a shorter completion time corresponding to better group performance. The robots were organized in a strict dominance hierarchy such that whenever two or more robots simultaneously had objects to deliver to the goal, the most dominant individual was allowed to proceed, while the less dominant individual(s) each waited their turn. The four robots, however, were not equally efficient at performing the task. The code for each robot was identical, except that the maximum speed was limited to different values, as follows: Robot0 "full-speed" (≈ 0.5 ft/sec); Robot1 "two-thirds-speed" (≈ 0.33 ft/sec); Robot2 "half-speed" (≈ 0.25 ft/sec); and Robot3 "one-third-speed" (≈ 0.17 ft/sec).

We conducted three sets of experiments, two with fixed hierarchies as baselines of comparison to the third, which allowed hierarchy restructuring through the use of AMMs. The experiments were designated as follows:

1. **Control**: The robots were members of a fixed hierarchy with the relative dominance of each inversely proportional to its maximum speed. Thus, Robot3 (the slowest) was the most dominant, and Robot0 (the fastest) was the least dominant.

2. **Optimal**: Complementary to Control, these experiments had the robots arranged in a fixed, optimal hierarchy, with the fastest as most dominant, and slowest as least dominant.

3. **Dynamic Leader Selection (DLS)**: The hierarchy was initialized to be identical to that of the Control experiments, but allowed hierarchy restructuring to improve performance.

In the DLS experiments, with no *a priori* information about a robot's speed provided, an AMM for each robot was constructed at run-time and used to evaluate performance. The metric of evaluation employed was a ratio giving the number of pucks per unit time that a robot is able to deliver: the higher this value, the faster the robot delivers pucks, and the better its performance. Each robot began a trial with its performance value initialized to zero. As it executed the task, its AMM was continuously updated, as was the performance value derived from it. The robot's position in the hierarchy was also updated so that it was more dominant than all other robots with lower performance values.

Table 1 presents the average time to completion (i.e., group performance) for the three experiments. In the experiments using dynamic leader selection we see a significant improvement in the time to completion over the Control experiments, mirroring a successful restructuring of the hierarchies to a more optimal configuration. The Optimal time is slightly, though not significantly, lower than the DLS time. This difference may be attributed to the fact that the DLS experiments are initially configured with the less efficient Control hierarchy structure.

	Control	DLS	Optimal
Mean time to completion	27.2	23.4	22.4
Standard deviation	1.1	1.3	1.1

Table 1: Mean time to completion for the Control, Dynamic Leader Selection, and Optimal experiments.

Published Papers:

Goldberg, Dani and Matarić, Maja J., "Mobile Robot Group Coordination Using a Model of Interaction Dynamics", *Proceedings of the SPIE: Sensor Fusion and Decentralized Control in Robotic Systems II*, Gerard T. McKee and Paul S. Schenker, eds., SPIE, 1999, 63-73.

Goldberg, Dani and Matarić, Maja J., "Coordinating Mobile Robot Group Behavior Using a Model of Interaction Dynamics", *Proceedings of Autonomous Agents '99*, Oren Etzioni, Jorg P. Juller, and Jeffrey M. Bradshaw, eds., ACM Press, 1999, 100-107.

7.4 AMMs for Regime Detection

In certain classes of mobile robot tasks, it may be necessary for a robot to detect significant global changes in the environment and modify its behavior or the task structure accordingly. The environment can be in a particular regime (i.e., a period of steady state) and then switch to a different regime requiring the robot to modify its behavior. Detecting such environmental regime changes may be difficult for a number of reasons:

- The robot may have no *a priori* knowledge of the environment and thus also lack a baseline for gauging environmental shifts.
- Given only local sensing capabilities, the robot may require a significant amount of time to estimate the state of the environment. Any estimate of state, however, may be outdated in a non-stationary system.
- The nature of the task may be stochastic, with uncertainties large enough to preclude an effective predictive model of environmental state, or dynamics too complex to make the development of such a model feasible or tractable. Alternatively, however potentially simple the system, there may be no *a priori* data with which to instantiate a model.
- Depending on the task or environment, the time scale of the environmental change that must be detected may differ. For example, in one task, the environmental change may be almost instantaneous, detectable between one moment and the next. In another task, the change may be slow and incremental, requiring the examination of a large time interval for detection. Hard-coding the robot with a specific time scale to use for regime detection can be problematic. A time scale that is too small makes the robot incapable of detecting the change. Conversely, a time scale that is unnecessarily large increases the time required to detect the change and may be undesirable in time-critical situations.

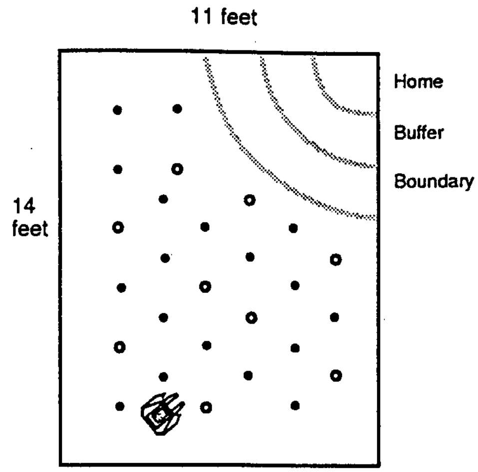


Figure 7: The land mine collection task. Open circles represent large mines and closed circles represent small mines. The robot can only deliver one mine at a time to the goal (Home).

As a concrete example, consider the task of collecting undetonated land mines in a field (a type of foraging). There are two types of mines, large and small, with destructive power proportional to their size (Figure 7). In this scenario, the robot is only able to carry one mine at a time, producing a large cost (in time) for each mine collected. It is important that the more destructive large mines be collected first, but that the robot be able to decide when to switch to the smaller mines. (Here we assume that the task requires the robot to collect one type of mine at a time. Alternatively, the robot might switch between types as necessary. We explore this alternative when we consider a reward maximization scenario in the next section.)

The difficulty of this task is compounded when the issues mentioned above apply. The robot may have no *a priori* information about the numbers of large and small mines in the field, their distributions, or relative proportions. The robot may also lack global sensing of the mines in the field and may not know the time scale appropriate to its decision for switching between mine types. This decision is dependent on factors including the size of the field and the relative densities of the two types of mines.

We have developed a mechanism for regime detection that resolves the above issues. The approach uses multiple *augmented Markov models* (AMMs). The AMMs are used to capture, in real time, the dynamics of a robot interacting with its environment in terms of the behaviors it performs. One AMM is created and maintained at each time scale that is monitored, and statistics about the environment at that time scale are derived from it. As task execution continues, AMMs are dynamically generated to accommodate the increasing time intervals. Sets of statistics from the models are used to determine whether the environmental regime has changed. This approach requires no *a priori* knowledge, uses only local sensing, and captures the notion of time scale. Additionally, it works naturally with stochastic task domains where variations between trials may change the most appropriate time scale for regime detection. We have validated the approach on a mobile robot performing the mine collection task.

Published Papers:

Goldberg, Dani and Matarić, Maja J., "Detecting Regime Changes with a Mobile Robot using Multiple Models", USC Institute for Robotics and Intelligent Systems Technical Report IRIS-00-382, 2000.

7.5 AMMs for Reward Maximization

In certain classes of mobile robot tasks, a robot may be required to perform optimally with respect to the information it possesses about the structure of its environment. Reward maximization may be used as a means of quantifying performance. In that framework, the robot receives reward (e.g., points) in proportion to its performance. Reward maximization in a non-stationary environment requires the robot to be able to estimate the state of the changing environment. There are a number of issues that can compound the difficulty of this problem. Some of them we have mentioned in the previous section:

- The robot may have no *a priori* knowledge of the environment.
- The robot may be limited to local sensing.
- The task may be stochastic.
- The time scale at which the non-stationarity of the environment manifests and thus can be detected may depend heavily on the task.

In addition to these issues, there is the further difficulty that in a stochastic system, the variability of execution must be considered in relation to detection of the non-stationarity. The variability associated with performing a task (or elements thereof) may be enormous and effectively mask gradual shifts in the environment. Conversely, in a system with very low variability, even minute shifts may be easily detected. Thus, effective estimation of environmental state requires an understanding of the system's variability (as often measured by variances, covariances, etc.).

Similar to our previous experimental scenario, consider the task of collecting undetonated land mines in a field. Assume that there are two types of mines, large and small, with destructive power proportional to their size. The robot's goal is to minimize the total destructive power of the mine field as much as possible during a given period of time. When the robot is given points in proportion to the destructive power of the mines it collects, the goal becomes equivalent to reward maximization. To accomplish its goal, the robot must have enough data about its environment (the field) to intelligently decide whether it is best to collect large mines or small ones at each point in time. The difficulty of this task is compounded when the issues mentioned above apply. The heart of this problem is to use the best possible estimate of environmental state given the limitations of the system.

We have developed an algorithm that provides a moving average estimate of the state of a non-stationary system. The algorithm dynamically adjusts the window size used in the moving average to accommodate the variances and type of non-stationarity exhibited by the system, while discarding outdated and redundant data. Multiple AMMs are learned, capturing in real time the dynamics of a robot interacting with its environment in terms of the behaviors it performs. One AMM is learned and maintained at each time scale that is monitored, and statistics about the environment at that time scale are derived from it. The state of the environment is thus estimated indirectly though the robot's interaction with it. As task execution continues, AMMs are dynamically generated to accommodate the increasing time intervals. Sets of statistics from the models are used to determine whether old data and AMMs are redundant/outdated and can be discarded. This approach requires no *a priori* knowledge, uses only local sensing, and captures the notion of time scale. Additionally, it works naturally with stochastic task domains where variations between trials may change the most appropriate amount of data for state estimation.

We have validated our approach using an implementation of the mine collection task with a real mobile robot and in simulation. We have conducted experiments in environments with both abruptly changing and gradually shifting non-stationarities. The data substantiate the effectiveness of our moving average algorithm using AMMs.

Published Papers:

Goldberg, Dani and Matarić, Maja J., "Learning Multiple Models for Reward Maximization", *Proceedings of the Seventeenth International Conference on Machine Learning*, Pat Langley, ed., Morgan Kaufman Publishers, 2000, 319-326.

Goldberg, Dani and Matarić, Maja J., "Reward Maximization in a Non-Stationary Mobile Robot Environment", *Proceedings of the Fourth International Conference on Autonomous Agents*, Charles Sierra, Maria Gini, and Jeffrey S. Rosenschein, eds., ACM Press, 2000, 92-99.

8 Accomplishments and Significance

The research enabled and supported by this ONR grant has produced scientific as well as practical results. As demonstrated by the very large number of publications, the scientific accomplishments have been recognized by the robotics community. Largely based on this work, the PI has received several awards for research:

- USC School of Engineering Junior Research Award 2000
- IEEE Robotics and Automation Society Early Career Award 2000
- MIT Technology Review TR100 Innovation Award 1999
- ACM Paper Award for co-authored student paper Agents-99
- NSF Career Award 1996-2000

Note that one of the above, the ACM Best Paper award, is for a paper, co-authored with a student funded by this grant, describing research on AMMs funded specifically by this grant.

This research has also had implications on teaching and the general public. The PI received the USC Innovative Undergraduate Teaching Award 1999-2000 for the class designed on the principles of robot control developed with this grant. Media attention to this research has been quite overwhelming. Some selected media coverage about the PI and this research, all unsolicited by the PI, is listed below.

- One of 7 scientists (including Nobel laureate Gertrude Elion, Ashok Gadgil, Michio Kaku, Steven Pinker, Karol Sikora, and Patricia Wright) featured in "Me & Isaac Newton", a film directed by Michael Apted, to be released in 2000.
- PBS Scientific American Frontiers "Natural Born Robots", hosted by Alan Alda, Nov 2, 1999.
- ABC World News Tonight with Peter Jennings, May 5, 1999.
- ABC Radio in Perth, Australia, Jul 23, 1998.
- BBC World Service in London, UK, Jul 21, 1998.
- The Washington Times, DC, 'Robotics convention stresses practicality" by Joann Loviglo, Aug 2, 1997.
- Mademoiselle, article on social behavior by Tonice Sgrignoli, Oct 1997.
- The Boston TAB, "Imagine This! From local labs and universities come 10 ideas that will change our lives", cover story by Courtney Claire Brigham, Jun 3, 1997.
- Wired, Japan, "Herd Mentality" by Jerry Shine, translated by S. Enami, May 1997.
- Electronics Times, "Ant approach aids Nerd Herd" by David Larner, Mar 6, 1997.
- Beyond 2000, hosted by Pat McGuinness, Nov 12, 1996.
- Computer Zeitung, Germany, by Ruth Henke and Rainer Scharf, Nov 7, 1996.
- Focus Magazine, UK, "Invasion of the Robots" by Sean Blair, Oct 1996.
- Discover Channel AI Series, interview by Cliff Lonsdale and Jane Hawkes, Sep 16, 1996.
- Wired, "Herd Mentality" by Jerry Shine, Jun 1996.
- Popular Science, "Go team Go!" by Steve Nadis with Jerry Shine, May 1996.
- New Zealand Public Radio, Mar 30, 1996.
- MIT Technology Review, article by Robert J. Crawford, Apr 1996.
- Utne Reader, "The Sharebots" by Carl Zimmer, Jan 1996.

The research described here has also been transitioned into Navy-relevant application areas. Specifically, we have worked closely with Christiane Duarte of NUWC and have helped her establish a Group Robotics Laboratory, with a heterogeneous group of various small wheeled robots networked via wireless Ethernet. We have held a workshop, with presentations by the PI and the students funded by this ONR grant (specifically Barry Werger and Dani Goldberg), to further help and inform the members of Duarte's team. At this time, members of her laboratory are already using Ayllu, the language/architecture in which BLE is implemented, developed in our lab, on their Pioneer robots, and adapting it to other platforms and simulators as the main behavior-coding and communication technology. BLE will therefore be used in their development and experiments, and plans include use of BLE to link AUVs as well. Similarly, our work towards the behavior-based Common Control Language will continue to be directly used on NUWC Group Robotics Lab Ayllu-based platforms.

We have also worked with Chris Duarte to apply the concepts of robot chaining to mine-sweeping operations at NUWC. A chain of robots maintaining sensor contact, sweeping in a circle, can potentially provide coverage guarantees and approximate locations of detected mines without need for localization or global communication capabilities. This is an immediate application being explored, although our results are being considered by other researchers (including those at Sandia National Laboratories), for mine-sweeping operations in different environments (land, surzone, etc.).

Web Dissemination

The PI has developed and maintains a large collection of informative Web pages on this research:
http://robotics.usc.edu/~maja/robot-control.html
http://robotics.usc.edu/~maja/bbs.html
http://robotics.usc.edu/~maja/gruop.html
http://robotics.usc.edu/~maja/learning.html
http://robotics.usc.edu/~barry/Chaining.html
http://robotics.usc.edu/ dani/hetero-homogeneous-groups.html
http://robotics.usc.edu/~barry/ullanta/UPRsoccer.html
http://robotics.usc.edu/~barry/BLE
http://robotics.usc.edu/~agents/projects/amms.html

List of Publications

Note: this list is a superset of the papers listed throughout this report, since not publications all were listed in specific categories above.

Refereed Journal Papers (10)

Werger, Barry B. and Matarić, Maja J., "From Insect to Internet: Situated Control for Networked Robot Teams", *Annals of Mathematics and Artificial Intelligence*, 2000.

Werger, Barry B. "Cooperation Without Deliberation: A Minimal Behavior-based Approach to Multi-robot Teams", *Artificial Intelligence*, 110, 1999, 293–320.

Michaud, François and Matarić, Maja J., "Representation of behavioral history for learning in nonstationary conditions", *Robotics and Autonomous Systems*, 29(2), Nov 30, 1999.

Fontan, Miguel S. and Matarić, Maja J., "Territorial Multi-Robot Task Division", *IEEE Transactions on Robotics and Automation*, 14(5), Oct 1998.

Matarić, Maja J., "Using Communication to Reduce Locality in Distributed Multi-Agent Learning", *Journal of Experimental and Theoretical Artificial Intelligence*, special issue on

Learning in DAI Systems, Gerhard Weiss, ed., 10(3), Jul-Sep, 1998, 357-369.

Matarić, Maja J., "Coordination and Learning in Multi-Robot Systems", IEEE Intelligent Systems, Mar/Apr 1998, 6-8.

Michaud, François and Matarić, Maja J., "Learning from History for Behavior-Based Mobile Robots in Non-Stationary Conditions", *Autonomous Robots*, 5(3-4), Jul/Aug 1998, 335-354, and *Machine Learning*, 31(1-3), 141-167, joint special issue on "Learning in Autonomous Robots."

Matarić, Maja J., "Behavior-Based Robotics as a Tool for Synthesis of Artificial Behavior and Analysis of Natural Behavior", *Trends in Cognitive Science*, 2(3), Mar 1998, 82-87.

Matarić, Maja J., "Behavior-Based Control: Examples from Navigation, Learning, and Group Behavior", *Journal of Theoretical and Experimental Artificial Intelligence*, special issue on Software Architectures for Physical Agents, 9(2-3), H. Hexmoor, I. Horswill, and D. Kortenkamp, eds., 1997, 323-336.

Matarić, Maja J., "Behavior-Based Robotics", invited contribution to the *MIT Encyclopedia of Cognitive Science*, R. Wilson and F. Keil, eds., MIT Press, April 1999, 74-77.

Refereed Conference Papers (13)

Werger, Barry B. and Matarić, Maja J., "Broadcast of Local Eligibility for Multi-Target Observation", *Distributed Autonomous Robotic Systems 4, Proceedings of DARS 2000*, Knoxville, Tennessee.

Duarte, Christiane N. and Werger, Barry B., "Defining a Common Control Language for Multiple Autonomous Vehicle Operations", *Proceedings of OCEANS 2000 MTS/IEEE*, Providence, Rhode Island, September, 2000.

Goldberg, Dani and Matarić, Maja J., "Learning Multiple Models for Reward Maximization", *Proceedings of the Seventeenth International Conference on Machine Learning*, Pat Langley, ed., Morgan Kaufman Publishers, 2000, 319-326.

Goldberg, Dani and Matarić, Maja J., "Reward Maximization in a Non-Stationary Mobile Robot Environment", *Proceedings of the Fourth International Conference on Autonomous Agents*, Charles Sierra, Maria Gini, and Jeffrey S. Rosenschein, eds., ACM Press, 2000, 92-99.

Goldberg, Dani and Matarić, Maja J., "Mobile Robot Group Coordination Using a Model of Interaction Dynamics", *Proceedings of the SPIE: Sensor Fusion and Decentralized Control in Robotic Systems II*, Gerard T. McKee and Paul S. Schenker, eds., SPIE, 1999, 63-73.

Goldberg, Dani and Matarić, Maja J., "Coordinating Mobile Robot Group Behavior Using a Model of Interaction Dynamics", *Proceedings of Autonomous Agents '99*, Oren Etzioni, Jorg

P. Juller, and Jeffrey M. Bradshaw, eds., ACM Press, 1999, 100-107.

Michaud, François and Matarić, Maja J., "Learning from History for Adaptive Mobile Robot Control", *Proceedings, IROS-98*, Victoria, BC, Canada, Oct 12-16, 1998.

Michaud, François and Matarić, Maja J., "A History-Based Approach for Adaptive Robot Behavior in Dynamic Environments", *Autonomous Agents '98*, Katia P. Sycara and Michael Wooldridge, eds., ACM Press, 1998, 422-429.

Matarić, Maja J., "Using Communication to Reduce Locality in Distributed Multi-Agent Learning", *Proceedings, AAAI-97*, Providence, RI, Jul 27-31, 1997, 643-648.

Goldberg, Dani and Matarić, Maja J., "Interference as a Tool for Designing and Evaluating Multi-Robot Controllers", *Proceedings of the Fourteenth National Conference on Artificial Intelligence*, AAAI Press, 1997.

Michaud, François and Matarić, Maja J., "Behavior Evaluation and Learning from an Internal Point of View", *Proceedings, FLAIRS-97*, Daytona, Florida, May 1997.

Fontan, Miguel S. and Matarić, Maja J., "A Study of Territoriality: The Role of Critical Mass in Adaptive Task Division", *Proceedings, From Animals to Animats 4, 4th International Conference on Simulation of Adaptive Behavior (SAB-96)*, P. Maes, M. Matarić, J-A. Meyer, J. Pollack, and S. Wilson, eds., MIT Press, 1996, 553-561.

Werger, Barry B. and Matarić, Maja J., "Robotic Food Chains: Externalization of State and Program for Minimal-Agent Foraging," *From Animals to Animats 4, Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, MIT Press, pp. 625-634.

Refereed Conference Posters (4)

Werger, Barry and Matarić, Maja J., "Broadcast of Local Eligibility: Behavior-Based Control for Strongly-Cooperative Robot Teams", *Proceedings of the Fourth International Conference on Autonomous Agents*, Charles Sierra, Maria Gini, and Jeffrey S. Rosenschein, eds., ACM Press, 2000, 21-22.

Sankaranarayanan, Aruna, S. and Matarić, Maja J., "The Multi-Agent-based Schedule Calculator (MASC) System", *Autonomous Agents '98*, Katia P. Sycara and Michael Wooldridge, eds., ACM Press, 1998, 465-466.

Werger, Barry B. and Matarić, Maja J., "Quick 'n' Dirty Generalization for Mobile Robot Learning", *IJCAI-97*, Nagoya, Japan, Aug 26-28, 1997.

Matarić, Maja J., "Studying the Role of Embodiment in Cognition", *Annual Meeting of the Society for Philosophy and Psychology*, The New School for Social Research, New York, Jun

5-8, 1997.

Technical Reports (8)

Goldberg, Dani and Matarić, Maja J., "Robust Behavior-Based Control for Distributed Multi-Robot Collection Tasks", USC Institute for Robotics and Intelligent Systems Technical Report IRIS-00-387, 2000. Also submitted to *IEEE Transactions on Robotics and Automation.*

Goldberg, Dani and Matarić, Maja J., "Detecting Regime Changes with a Mobile Robot using Multiple Models", USC Institute for Robotics and Intelligent Systems Technical Report IRIS-00-382, 2000.

Werger, Barry B. and Matarić, Maja J. "Exploiting embodiment in multi-robot teams", USC Institute for Robotics and Intelligent Systems Technical Report IRIS-99-378, 1999.

Goldberg, Dani and Matarić, Maja J., "Augmented Markov Models", USC Institute for Robotics and Intelligent Systems Technical Report IRIS-99-367, 1999.

Michaud, François and Matarić, Maja J., "A History-Based Learning Approach for Adaptive Robot Behavior Selection", *Brandeis University Computer Science Technical Report* CS-97-192, Jul 1997.

Goldberg, Dani and Matarić, Maja J., "Interference as a Guide for Designing Efficient Group Behaviors", Brandeis University Computer Science Technical Report CS-96-186, 1996.

Matarić, Maja J., "Using Communication to Reduce Locality in Distributed Multi-Agent Learning", *Brandeis University Computer Science Technical Report* CS-96-190, Nov 1996.

Fontan, Miguel S. and Matarić, Maja J., "The Role of Critical Mass in Multi-Robot Adaptive Task Division", *Brandeis University Computer Science Technical Report* CS-96-187, Oct 1996.

Symposia and Workshops (2)

Tambe, Milind, Shen, Wei-min, Matarić, Maja J., Pynadath, David, Goldberg, Dani, Modi, Jay, Qiu, Zhun, Salemi, Behnam, "Team Work in Cyberspace: Using TEAMCORE to Make Agents Team-Ready", *Proceedings of the 1999 AAAI Spring Symposium.*

Matarić, Maja J., "Studying the Role of Embodiment in Cognition", *AAAI Fall Symposium on Embodied Cognition and Action*, MIT, Cambridge, MA, Nov 9-11, 1996.

References

Brooks, R. A. (1990*a*), Challenges for Complete Creature Architectures, *in* 'Proceedings of the 1st International Conference on Simulation of Adaptive Behavior'.

Brooks, R. A. (1990*b*), 'Elephants don't play chess', *Robotics and Autonomous Systems* **6**, 3–16.

Brooks, R. A. (1991*a*), Intelligence without reason, *in* 'Proceedings of IJCAI', Sydney, Australia, pp. 569–595.

Brooks, R. A. (1991*b*), 'Intelligence without representation', *Artificial Intelligence* **47**, 139–159.

Gordon, D. M. (1999), *Ants at Work*, The Free Press, New York.

Matarić, M. (1995), 'Issues and Approaches in the Design of Collective Autonomous Agents', *Robotics and Autonomous Systems.*

Matarić, M. J. (1995a), 'Designing and Understanding Adaptive Group Behavior', *Adaptive Behavior* **4**(1), 50–81.

Matarić, M. J. (1997), 'Behavior-Based Control: Examples from Navigation, Learning, and Group Behavior', *Journal of Experimental and Theoretical Artificial Intelligence* **9**(2–3), 323–336.

McFarland, D. (1987), *The Oxford Companion to Animal Behavior*, Oxford University Press.

Michaud, F. & Matarić, M. J. (1997), Behavior Evaluation and Learning from an Internal Point of View, *in* 'Proceedings, FLAIRS-97', Daytona, FL.

Michaud, F. & Matarić, M. J. (1998*a*), A History-Based Approach for Adaptive Robot Behavior in Dynamic Environments, *in* 'Autonomous Agents', ACM Press, Minneapolis, St. Paul, MI, pp. 422–429.

Michaud, F. & Matarić, M. J. (1998*b*), Learning From History for Adaptive Multi-Robot Control, *in* 'IROS–98', Victoria, BC, Canada.

Michaud, F. & Matarić, M. J. (1998*c*), 'Learning from History for Behavior-Based Mobile Robots in Non-stationary Conditions', *Autonomous Robots and Machine Learning* **AR:5,ML:31**, AR:335–354,ML:141–167.

Steels, L. (1994), 'The Artificial Life Roots of Artificial Intelligence', *Artificial Life.*

Werger, B. B. (1999), 'Cooperation Without Deliberation: A minimal behavior-based approach to multi-robot teams', *Artificial Intelligence* **110**, 293–320.

Werger, B. B. & Matarić, M. J. (2000), 'From Insect to Internet: Situated Control for Networked Robot Teams', *Annals of Mathematics and Artificial Intelligence.*